

## 5800-79B SEQLIST.TXT

## SEQUENCE LISTING

<110> Glucksman, Maria Alexandra  
 williamson, Mark  
 Tsia, Fong-Ying  
 Rudolph-Owen, Laura A.

<120> 22438, 23553, 25278, and 26212 Novel  
 Human Sulfatases (A CIP Application)

<130> 35800/208398(5800-79

<150> US 09/495,823

<151> 2000-01-31

<160> 14

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 525

<212> PRT

<213> homo sapiens

<400> 1

Met	Gly	Trp	Leu	Phe	Leu	Lys	Val	Leu	Leu	Ala	Gly	Val	Ser	Phe	Ser
1				5					10					15	
Gly	Phe	Leu	Tyr	Pro	Leu	Val	Asp	Phe	Cys	Ile	Ser	Gly	Lys	Thr	Arg
		20					25					30			
Gly	Gln	Lys	Pro	Asn	Phe	Val	Ile	Ile	Leu	Ala	Asp	Asp	Met	Gly	Trp
		35				40					45				
Gly	Asp	Leu	Gly	Ala	Asn	Trp	Ala	Glu	Thr	Lys	Asp	Thr	Ala	Asn	Leu
	50				55				60						
Asp	Lys	Met	Ala	Ser	Glu	Gly	Met	Arg	Phe	Val	Asp	Phe	His	Ala	Ala
65					70				75					80	
Ala	Ser	Thr	Cys	Ser	Pro	Ser	Arg	Ala	Ser	Leu	Leu	Thr	Gly	Arg	Leu
			85					90					95		
Gly	Leu	Arg	Asn	Gly	Val	Thr	Arg	Asn	Phe	Ala	Val	Thr	Ser	Val	Gly
			100					105					110		
Gly	Leu	Pro	Leu	Asn	Glu	Thr	Thr	Leu	Ala	Glu	Val	Leu	Gln	Gln	Ala
		115				120					125				
Gly	Tyr	Val	Thr	Gly	Ile	Ile	Gly	Lys	Trp	His	Leu	Gly	His	His	Gly
	130				135						140				
Ser	Tyr	His	Pro	Asn	Phe	Arg	Gly	Phe	Asp	Tyr	Tyr	Phe	Gly	Ile	Pro
145					150				155					160	
Tyr	Ser	His	Asp	Met	Gly	Cys	Thr	Asp	Thr	Pro	Gly	Tyr	Asn	His	Pro
			165					170					175		
Pro	Cys	Pro	Ala	Cys	Pro	Gln	Gly	Asp	Gly	Pro	Ser	Arg	Asn	Leu	Gln
			180					185					190		
Arg	Asp	Cys	Tyr	Thr	Asp	Val	Ala	Leu	Pro	Leu	Tyr	Glu	Asn	Leu	Asn
		195					200					205			
Ile	Val	Glu	Gln	Pro	Val	Asn	Leu	Ser	Ser	Leu	Ala	Gln	Lys	Tyr	Ala
	210					215					220				
Glu	Lys	Ala	Thr	Gln	Phe	Ile	Gln	Arg	Ala	Ser	Thr	Ser	Gly	Arg	Pro
225					230				235					240	
Phe	Leu	Leu	Tyr	Val	Ala	Leu	Ala	His	Met	His	Val	Pro	Leu	Pro	Val
			245					250					255		
Thr	Gln	Leu	Pro	Ala	Ala	Pro	Arg	Gly	Arg	Ser	Leu	Tyr	Gly	Ala	Gly
		260					265						270		
Leu	Trp	Glu	Met	Asp	Ser	Leu	Val	Gly	Gln	Ile	Lys	Asp	Lys	Val	Asp
		275					280					285			

## 5800-79B SEQLIST.TXT

His Thr Val Lys Glu Asn Thr Phe Leu Trp Phe Thr Gly Asp Asn Gly  
 290 295 300  
 Pro Trp Ala Gln Lys Cys Glu Leu Ala Gly Ser Val Gly Pro Phe Thr  
 305 310 315 320  
 Gly Phe Trp Gln Thr Arg Gln Gly Gly Ser Pro Ala Lys Gln Thr Thr  
 325 330 335  
 Trp Glu Gly Gly His Arg Val Pro Ala Leu Ala Tyr Trp Pro Gly Arg  
 340 345 350  
 Val Pro Val Asn Val Thr Ser Thr Ala Leu Leu Ser Val Leu Asp Ile  
 355 360 365  
 Phe Pro Thr Val Val Ala Leu Ala Gln Ala Ser Leu Pro Gln Gly Arg  
 370 375 380  
 Arg Phe Asp Gly Val Asp Val Ser Glu Val Leu Phe Gly Arg Ser Gln  
 385 390 395 400  
 Pro Gly His Arg Val Leu Phe His Pro Asn Ser Gly Ala Ala Gly Glu  
 405 410 415  
 Phe Gly Ala Leu Gln Thr Val Arg Leu Glu Arg Tyr Lys Ala Phe Tyr  
 420 425 430  
 Ile Thr Gly Gly Ala Arg Ala Cys Asp Gly Ser Thr Gly Pro Glu Leu  
 435 440 445  
 Gln His Lys Phe Pro Leu Ile Phe Asn Leu Glu Asp Asp Thr Ala Glu  
 450 455 460  
 Ala Val Pro Leu Glu Arg Gly Gly Ala Glu Tyr Gln Ala Val Leu Pro  
 465 470 475 480  
 Glu Val Arg Lys Val Leu Ala Asp Val Leu Gln Asp Ile Ala Asn Asp  
 485 490 495  
 Asn Ile Ser Ser Ala Asp Tyr Thr Gln Asp Pro Ser Val Thr Pro Cys  
 500 505 510  
 Cys Asn Pro Tyr Gln Ile Ala Cys Arg Cys Gln Ala Ala  
 515 520 525

<210> 2  
 <211> 2175  
 <212> DNA  
 <213> homo sapiens

<220>  
 <221> CDS  
 <222> (248)...(1825)

<400> 2  
 cacgcgtccg caaatttcct gattcttttg aattaggatt ccagatgggg gcctcatttc 60  
 tacagccccc aacattccta tagccgttat cactgccatc accactgcca ccagcatctt 120  
 cttgcagatt ccaccctgc tccccagaga cttcctgctt tgaaagtgag cagaaaggaa 180  
 gctctcagaa aaatctctag tgggtgctgc cgctgcctcca gacaatcga atcctgcctt 240  
 caccacc atg ggc tgg ctt ttt cta aag gtt ttg ttg gcg gga gtg agt 289  
 Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser  
 1 5 10  
  
 ttc tca gga ttt ctt tat cct ctt gtg gat ttt tgc atc agt ggg aaa 337  
 Phe Ser Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys  
 15 20 25 30  
  
 aca aga gga cag aag cca aac ttt gtg att att ttg gcc gat gac atg 385  
 Thr Arg Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met  
 35 40 45  
  
 ggg tgg ggt gac ctg gga gca aac tgg gca gaa aca aag gac act gcc 433  
 Gly Trp Gly Asp Leu Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala  
 50 55 60  
  
 aac ctt gat aag atg gct tcg gag gga atg agg ttt gtg gat ttc cat 481  
 Page 2

5800-79B SEQLIST.TXT

Asn	Leu	Asp 65	Lys	Met	Ala	Ser	Glu 70	Gly	Met	Arg	Phe	Val 75	Asp	Phe	His	
gca Ala	gct Ala 80	gcc Ala	tcc Ser	acc Thr	tgc Cys	tca Ser 85	ccc Pro	tcc Ser	cgg Arg	gct Ala	tcc Ser 90	ttg Leu	ctc Leu	acc Thr	ggc Gly	529
cgg Arg 95	ctt Leu	ggc Gly	ctt Leu	cgc Arg	aat Asn 100	gga Gly	gtc Val	aca Thr	cgc Arg	aac Asn 105	ttt Phe	gca Ala	gtc Val	act Thr	tct Ser 110	577
gtg Val	gga Gly	ggc Gly	ctt Leu	ccg Pro 115	ctc Leu	aac Asn	gag Glu	acc Thr	acc Thr 120	ttg Leu	gca Ala	gag Glu	gtg Val	ctg Leu 125	cag Gln	625
cag Gln	gcg Ala	ggt Gly	tac Tyr 130	gtc Val	act Thr	ggg Gly	ata Ile	ata Ile 135	ggc Gly	aaa Lys	tgg Trp	cat His	ctt Leu 140	gga Gly	cac His	673
cac His	ggc Gly	tct Ser 145	tat Tyr	cac His	ccc Pro	aac Asn	ttc Phe 150	cgt Arg	ggt Gly	ttt Phe	gat Asp	tac Tyr 155	tac Tyr	ttt Phe	gga Gly	721
atc Ile	cca Pro 160	tat Tyr	agc Ser	cat His	gat Asp	atg Met 165	ggc Gly	tgt Cys	act Thr	gat Asp	act Thr 170	cca Pro	ggc Gly	tac Tyr	aac Asn	769
cac His 175	cct Pro	cct Pro	tgt Cys	cca Pro	gcg Ala 180	tgt Cys	cca Pro	cag Gln	ggt Gly	gat Asp 185	gga Gly	cca Pro	tca Ser	agg Arg	aac Asn 190	817
ctt Leu	caa Gln	aga Arg	gac Asp	tgt Cys 195	tac Tyr	act Thr	gac Asp	gtg Val	gcc Ala 200	ctc Leu	cct Pro	ctt Leu	tat Tyr	gaa Glu 205	aac Asn	865
ctc Leu	aac Asn	att Ile	gtg Val 210	gag Glu	cag Gln	ccg Pro	gtg Val	aac Asn 215	ttg Leu	agc Ser	agc Ser	ctt Leu	gcc Ala 220	cag Gln	aag Lys	913
tat Tyr	gct Ala	gag Glu 225	aaa Lys	gca Ala	acc Thr	cag Gln	ttc Phe 230	atc Ile	cag Gln	cgt Arg	gca Ala	agc Ser 235	acc Thr	agc Ser	ggg Gly	961
agg Arg	ccc Pro 240	ttc Phe	ctg Leu	ctc Leu	tat Tyr	gtg Val 245	gct Ala	ctg Leu	gcc Ala	cac His	atg Met 250	cac His	gtg Val	ccc Pro	tta Leu	1009
ccc Pro 255	gtg Val	act Thr	cag Gln	cta Leu	cca Pro 260	gca Ala	gcg Ala	cca Pro	cgg Arg	ggc Gly 265	aga Arg	agc Ser	ctg Leu	tat Tyr	ggt Gly 270	1057
gca Ala	ggg Gly	ctc Leu	tgg Trp	gag Glu 275	atg Met	gac Asp	agt Ser	ctg Leu	gtg Val 280	ggc Gly	cag Gln	atc Ile	aag Lys	gac Asp 285	aaa Lys	1105
gtt Val	gac Asp	cac His	aca Thr 290	gtg Val	aag Lys	gaa Glu	aac Asn	aca Thr 295	ttc Phe	ctc Leu	tgg Trp	ttt Phe	aca Thr 300	gga Gly	gac Asp	1153
aat Asn	ggc Gly	ccg Pro 305	tgg Trp	gct Ala	cag Gln	aag Lys	tgt Cys 310	gag Glu	cta Leu	gcg Ala	ggc Gly	agt Ser 315	gtg Val	ggt Gly	ccc Pro	1201

## 5800-79B SEQLIST.TXT

ttc act gga ttt tgg caa act cgt caa ggg gga agt cca gcc aag cag Phe Thr Gly Phe Trp Gln Thr Arg Gln Gly Gly Ser Pro Ala Lys Gln 320 325 330	1249
acg acc tgg gaa gga ggg cac cgg gtc cca gca ctg gct tac tgg cct Thr Thr Trp Glu Gly Gly His Arg Val Pro Ala Leu Ala Tyr Trp Pro 335 340 345 350	1297
ggc aga gtt cca gtt aat gtc acc agc act gcc ttg tta agc gtg ctg Gly Arg Val Pro Val Asn Val Thr Ser Thr Ala Leu Leu Ser Val Leu 355 360 365	1345
gac att ttt cca act gtg gta gcc ctg gcc cag gcc agc tta cct caa Asp Ile Phe Pro Thr Val Val Ala Leu Ala Gln Ala Ser Leu Pro Gln 370 375 380	1393
gga cgg cgc ttt gat ggt gtg gac gtc tcc gag gtg ctc ttt ggc cgg Gly Arg Arg Phe Asp Gly Val Asp Val Ser Glu Val Leu Phe Gly Arg 385 390 395	1441
tca cag cct ggg cac agg gtg ctg ttc cac ccc aac agc ggg gca gct Ser Gln Pro Gly His Arg Val Leu Phe His Pro Asn Ser Gly Ala Ala 400 405 410	1489
gga gag ttt gga gcc ctg cag act gtc cgc ctg gag cgt tac aag gcc Gly Glu Phe Gly Ala Leu Gln Thr Val Arg Leu Glu Arg Tyr Lys Ala 415 420 425 430	1537
ttc tac att acc ggt gga gcc agg gcg tgt gat ggg agc acg ggg cct Phe Tyr Ile Thr Gly Gly Ala Arg Ala Cys Asp Gly Ser Thr Gly Pro 435 440 445	1585
gag ctg cag cat aag ttt cct ctg att ttc aac ctg gaa gac gat acc Glu Leu Gln His Lys Phe Pro Leu Ile Phe Asn Leu Glu Asp Asp Thr 450 455 460	1633
gca gaa gct gtg ccc cta gaa aga ggt ggt gcg gag tac cag gct gtg Ala Glu Ala Val Pro Leu Glu Arg Gly Gly Ala Glu Tyr Gln Ala Val 465 470 475	1681
ctg ccc gag gtc aga aag gtt ctt gca gac gtc ctc caa gac att gcc Leu Pro Glu Val Arg Lys Val Leu Ala Asp Val Leu Gln Asp Ile Ala 480 485 490	1729
aac gac aac atc tcc agc gca gat tac act cag gac cct tca gta act Asn Asp Asn Ile Ser Ser Ala Asp Tyr Thr Gln Asp Pro Ser Val Thr 495 500 505 510	1777
ccc tgc tgt aat ccc tac caa att gcc tgc cgc tgt caa gcc gca taa Pro Cys Cys Asn Pro Tyr Gln Ile Ala Cys Arg Cys Gln Ala Ala * 515 520 525	1825
cagaccaatt tttattccac gaggaggagt acctggaaat taggcaagtt tgcttccaaa tttcatTTTT accctcttta caaacacacg ctttagttta gtcttgagggt ttagtttttg agttagcctt gcatatccct tctgtatcct gtccctcctc cacgccgacc cgagagcagc tgagctgcgc tggctctggg caggagtggt gccttaattgg gaagcacacg ggctttggag tcaggcacag gtgccagctc cagcttttga acttgggcaa ttgtttaacc taacctgcaa gttgattttg aggggttaaata aaaggcatatc atgaaaaaaa aaaaaaaaaa	1885 1945 2005 2065 2125 2175

<210> 3  
 <211> 871  
 <212> PRT  
 <213> homo sapiens

## 5800-79B SEQLIST.TXT

<400> 3  
 Met Lys Tyr Ser Cys Cys Ala Leu Val Leu Ala Val Leu Gly Thr Glu  
 1 5 10 15  
 Leu Leu Gly Ser Leu Cys Ser Thr Val Arg Ser Pro Arg Phe Arg Gly  
 20 25 30  
 Arg Ile Gln Gln Glu Arg Lys Asn Ile Arg Pro Asn Ile Ile Leu Val  
 35 40 45  
 Leu Thr Asp Asp Gln Asp Val Glu Leu Gly Ser Leu Gln Val Met Asn  
 50 55 60  
 Lys Thr Arg Lys Ile Met Glu His Gly Gly Ala Thr Phe Ile Asn Ala  
 65 70 75 80  
 Phe Val Thr Thr Pro Met Cys Cys Pro Ser Arg Ser Ser Met Leu Thr  
 85 90 95  
 Gly Lys Tyr Val His Asn His Asn Val Tyr Thr Asn Asn Glu Asn Cys  
 100 105 110  
 Ser Ser Pro Ser Trp Gln Ala Met His Glu Pro Arg Thr Phe Ala Val  
 115 120 125  
 Tyr Leu Asn Asn Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr Leu  
 130 135 140  
 Asn Glu Tyr Asn Gly Ser Tyr Ile Pro Pro Gly Trp Arg Glu Trp Leu  
 145 150 155 160  
 Gly Leu Ile Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Val Cys Arg Asn  
 165 170 175  
 Gly Ile Lys Glu Lys His Gly Phe Asp Tyr Ala Lys Asp Tyr Phe Thr  
 180 185 190  
 Asp Leu Ile Thr Asn Glu Ser Ile Asn Tyr Phe Lys Met Ser Lys Arg  
 195 200 205  
 Met Tyr Pro His Arg Pro Val Met Met Val Ile Ser His Ala Ala Pro  
 210 215 220  
 His Gly Pro Glu Asp Ser Ala Pro Gln Phe Ser Lys Leu Tyr Pro Asn  
 225 230 235 240  
 Ala Ser Gln His Ile Thr Pro Ser Tyr Asn Tyr Ala Pro Asn Met Asp  
 245 250 255  
 Lys His Trp Ile Met Gln Tyr Thr Gly Pro Met Leu Pro Ile His Met  
 260 265 270  
 Glu Phe Thr Asn Ile Leu Gln Arg Lys Arg Leu Gln Thr Leu Met Ser  
 275 280 285  
 Val Asp Asp Ser Val Glu Arg Leu Tyr Asn Met Leu Val Glu Thr Gly  
 290 295 300  
 Glu Leu Glu Asn Thr Tyr Ile Ile Tyr Thr Ala Asp His Gly Tyr His  
 305 310 315 320  
 Ile Gly Gln Phe Gly Leu Val Lys Gly Lys Ser Met Pro Tyr Asp Phe  
 325 330 335  
 Asp Ile Arg Val Pro Phe Phe Ile Arg Gly Pro Ser Val Glu Pro Gly  
 340 345 350  
 Ser Ile Val Pro Gln Ile Val Leu Asn Ile Asp Leu Ala Pro Thr Ile  
 355 360 365  
 Leu Asp Ile Ala Gly Leu Asp Thr Pro Pro Asp Val Asp Gly Lys Ser  
 370 375 380  
 Val Leu Lys Leu Leu Asp Pro Glu Lys Pro Gly Asn Arg Phe Arg Thr  
 385 390 395 400  
 Asn Lys Lys Ala Lys Ile Trp Arg Asp Thr Phe Leu Val Glu Arg Gly  
 405 410 415  
 Lys Phe Leu Arg Lys Lys Glu Glu Ser Ser Lys Asn Ile Gln Gln Ser  
 420 425 430  
 Asn His Leu Pro Lys Tyr Glu Arg Val Lys Glu Leu Cys Gln Gln Ala  
 435 440 445  
 Arg Tyr Gln Thr Ala Cys Glu Gln Pro Gly Gln Lys Trp Gln Cys Ile  
 450 455 460  
 Glu Asp Thr Ser Gly Lys Leu Arg Ile His Lys Cys Lys Gly Pro Ser  
 465 470 475 480  
 Asp Leu Leu Thr Val Arg Gln Ser Thr Arg Asn Leu Tyr Ala Arg Gly  
 Page 5

## 490

```
<210> 4
<211> 4321
<212> DNA
<213> homo sapiens
```

```
<221> misc_feature
<222> 4310
<223> n = A,T,C or G
```

## 5800-79B SEQLIST.TXT

&lt;400&gt; 4

```

cccacgcgtc cggctaata atcttggggc cgggtgtcggg ccgggggcggc ttgatcggca 60
actaggaaac cccaggcgca gaggccagga gcgagggcag cgaggatcag aggccaggcc 120
ttcccggctg ccggcgctcc tcggagggtca gggcagatga ggaacatgac tctccccctt 180
cggaggagga aggaagtccc gctgccacct tatctctgct cctctgcctc ctccctgttc 240
ccagagcttt ttctctagag aagattttga aggcggcttt tgtgctgacg gccaccacc 300
atcatctaaa gaagataaac ttggcaaata acatgcagggt tcttcaaggc agaataattg 360
cagaaaatct tcaaaggacc ctatctgcag atgttctgaa tacctctgag aatagagatt 420
gattattcaa ccaggatacc taattcaaga actccagaaa tcaggagacg gagacatttt 480
gtcagttttg caacattgga ccaaataca atg aag tat tct tgc tgt gct ctg 533
                               Met Lys Tyr Ser Cys Cys Ala Leu
                               1                               5

gtt ttg gct gtc ctg ggc aca gaa ttg ctg gga agc ctc tgt tcg act 581
Val Leu Ala Val Leu Gly Thr Glu Leu Leu Gly Ser Leu Cys Ser Thr
    10                               15                               20

gtc aga tcc ccg agg ttc aga gga cgg ata cag cag gaa cga aaa aac 629
Val Arg Ser Pro Arg Phe Arg Gly Arg Ile Gln Gln Glu Arg Lys Asn
    25                               30                               35                               40

atc cga ccc aac att att ctt gtg ctt acc gat gat caa gat gtg gag 677
Ile Arg Pro Asn Ile Ile Leu Val Leu Thr Asp Asp Gln Asp Val Glu
                               45                               50                               55

ctg ggg tcc ctg caa gtc atg aac aaa acg aga aag att atg gaa cat 725
Leu Gly Ser Leu Gln Val Met Asn Lys Thr Arg Lys Ile Met Glu His
                               60                               65                               70

ggg ggg gcc acc ttc atc aat gcc ttt gtg act aca ccc atg tgc tgc 773
Gly Gly Ala Thr Phe Ile Asn Ala Phe Val Thr Thr Pro Met Cys Cys
    75                               80                               85

ccg tca cgg tcc tcc atg ctc acc ggg aag tat gtg cac aat cac aat 821
Pro Ser Arg Ser Ser Met Leu Thr Gly Lys Tyr Val His Asn His Asn
    90                               95                               100

gtc tac acc aac aac gag aac tgc tct tcc ccc tcg tgg cag gcc atg 869
Val Tyr Thr Asn Asn Glu Asn Cys Ser Ser Pro Ser Trp Gln Ala Met
    105                               110                               115                               120

cat gag cct cgg act ttt gct gta tat ctt aac aac act ggc tac aga 917
His Glu Pro Arg Thr Phe Ala Val Tyr Leu Asn Asn Thr Gly Tyr Arg
                               125                               130                               135

aca gcc ttt ttt gga aaa tac ctc aat gaa tat aat ggc agc tac atc 965
Thr Ala Phe Phe Gly Lys Tyr Leu Asn Glu Tyr Asn Gly Ser Tyr Ile
                               140                               145                               150

ccc cct ggg tgg cga gaa tgg ctt gga tta atc aag aat tct cgc ttc 1013
Pro Pro Gly Trp Arg Glu Trp Leu Gly Leu Ile Lys Asn Ser Arg Phe
    155                               160                               165

tat aat tac act gtt tgt cgc aat ggc atc aaa gaa aag cat gga ttt 1061
Tyr Asn Tyr Thr Val Cys Arg Asn Gly Ile Lys Glu Lys His Gly Phe
    170                               175                               180

gat tat gca aag gac tac ttc aca gac tta atc act aac gag agc att 1109
Asp Tyr Ala Lys Asp Tyr Phe Thr Asp Leu Ile Thr Asn Glu Ser Ile
    185                               190                               195                               200

aat tac ttc aaa atg tct aag aga atg tat ccc cat agg ccc gtt atg 1157

```

5800-79B SEQLIST.TXT

Asn	Tyr	Phe	Lys	Met 205	Ser	Lys	Arg	Met	Tyr 210	Pro	His	Arg	Pro	Val 215	Met	
atg Met	gtg Val	atc Ile	agc Ser 220	cac His	gct Ala	gcg Ala	ccc Pro	cac His 225	ggc Gly	ccc Pro	gag Glu	gac Asp	tca Ser 230	gcc Ala	cca Pro	1205
cag Gln	ttt Phe	tct Ser 235	aaa Lys	ctg Leu	tac Tyr	ccc Pro	aat Asn 240	gct Ala	tcc Ser	caa Gln	cac His	ata Ile 245	act Thr	cct Pro	agt Ser	1253
tat Tyr	aac Asn 250	tat Tyr	gca Ala	cca Pro	aat Asn 255	atg Met 255	gat Asp	aaa Lys	cac His	tgg Trp	att Ile 260	atg Met	cag Gln	tac Tyr	aca Thr	1301
gga Gly 265	cca Pro	atg Met	ctg Leu	ccc Pro	atc Ile 270	cac His	atg Met	gaa Glu	ttt Phe	aca Thr 275	aac Asn	att Ile	cta Leu	cag Gln	cgc Arg 280	1349
aaa Lys	agg Arg	ctc Leu	cag Gln	act Thr 285	ttg Leu	atg Met	tca Ser	gtg Val	gat Asp 290	gat Asp	tct Ser	gtg Val	gag Glu	agg Arg 295	ctg Leu	1397
tat Tyr	aac Asn	atg Met	ctc Leu 300	gtg Val	gag Glu	acg Thr	ggg Gly	gag Glu 305	ctg Leu	gag Glu	aat Asn	act Thr	tac Tyr 310	atc Ile	att Ile	1445
tac Tyr	acc Thr	gcc Ala 315	gac Asp	cat His	ggt Gly	tac Tyr	cat His 320	att Ile	ggg Gly	cag Gln	ttt Phe	gga Gly 325	ctg Leu	gtc Val	aag Lys	1493
ggg Gly 330	aaa Lys	tcc Ser	atg Met	cca Pro	tat Tyr	gac Asp 335	ttt Phe	gat Asp	att Ile	cgt Arg	gtg Val 340	cct Pro	ttt Phe	ttt Phe	att Ile	1541
cgt Arg 345	ggt Gly	cca Pro	agt Ser	gta Val	gaa Glu 350	cca Pro	gga Gly	tca Ser	ata Ile	gtc Val 355	cca Pro	cag Gln	atc Ile	gtt Val	ctc Leu 360	1589
aac Asn	att Ile	gac Asp	ttg Leu	gcc Ala 365	ccc Pro	acg Thr	atc Ile	ctg Leu	gat Asp 370	att Ile	gct Ala	ggg Gly	ctc Leu	gac Asp 375	aca Thr	1637
cct Pro	cct Pro	gat Asp	gtg Val 380	gac Asp	ggc Gly	aag Lys	tct Ser	gtc Val 385	ctc Leu	aaa Lys	ctt Leu	ctg Leu	gac Asp 390	cca Pro	gaa Glu	1685
aag Lys	cca Pro	ggt Gly 395	aac Asn	agg Arg	ttt Phe	cga Arg	aca Thr 400	aac Asn	aag Lys	aag Lys	gcc Ala 405	aaa Lys	att Ile	tgg Trp	cgt Arg	1733
gat Asp	aca Thr 410	ttc Phe	cta Leu	gtg Val	gaa Glu	aga Arg 415	ggc Gly	aaa Lys	ttt Phe	cta Leu	cgt Arg 420	aag Lys	aag Lys	gaa Glu	gaa Glu	1781
tcc Ser 425	agc Ser	aag Lys	aat Asn	atc Ile	caa Gln 430	cag Gln	tca Ser	aat Asn	cac His	ttg Leu 435	ccc Pro	aaa Lys	tat Tyr	gaa Glu	cgg Arg 440	1829
gtc Val	aaa Lys	gaa Glu	cta Leu	tgc Cys 445	cag Gln	cag Gln	gcc Ala	agg Arg	tac Tyr 450	cag Gln	aca Thr	gcc Ala	tgt Cys	gaa Glu 455	caa Gln	1877



5800-79B SEQLIST.TXT

ccg	ggg	cag	aag	tgg	caa	tgc	att	gag	gat	aca	tct	ggc	aag	ctt	cga	1925
Pro	Gly	Gln	Lys	Trp	Gln	Cys	Ile	Glu	Asp	Thr	Ser	Gly	Lys	Leu	Arg	
			460					465					470			
att	cac	aag	tgt	aaa	gga	ccc	agt	gac	ctg	ctc	aca	gtc	cgg	cag	agc	1973
Ile	His	Lys	Cys	Lys	Gly	Pro	Ser	Asp	Leu	Leu	Thr	Val	Arg	Gln	Ser	
		475					480					485				
acg	cgg	aac	ctc	tac	gct	cgc	ggc	ttc	cat	gac	aaa	gac	aaa	gag	tgc	2021
Thr	Arg	Asn	Leu	Tyr	Ala	Arg	Gly	Phe	His	Asp	Lys	Asp	Lys	Glu	Cys	
	490					495					500					
agt	tgt	agg	gag	tct	ggt	tac	cgt	gcc	agc	aga	agc	caa	aga	aag	agt	2069
Ser	Cys	Arg	Glu	Ser	Gly	Tyr	Arg	Ala	Ser	Arg	Ser	Gln	Arg	Lys	Ser	
505					510				515						520	
caa	cgg	caa	ttc	ttg	aga	aac	cag	ggg	act	cca	aag	tac	aag	ccc	aga	2117
Gln	Arg	Gln	Phe	Leu	Arg	Asn	Gln	Gly	Thr	Pro	Lys	Tyr	Lys	Pro	Arg	
				525					530					535		
ttt	gtc	cat	act	cgg	cag	aca	cgt	tcc	ttg	tcc	gtc	gaa	ttt	gaa	ggc	2165
Phe	Val	His	Thr	Arg	Gln	Thr	Arg	Ser	Leu	Ser	Val	Glu	Phe	Glu	Gly	
			540					545					550			
gaa	ata	tat	gac	ata	aat	ctg	gaa	gaa	gaa	gaa	gaa	ttg	caa	gtg	ttg	2213
Glu	Ile	Tyr	Asp	Ile	Asn	Leu	Glu	Glu	Glu	Glu	Glu	Leu	Gln	Val	Leu	
		555					560					565				
caa	cca	aga	aac	att	gct	aag	cgt	cat	gat	gaa	ggc	cac	aag	ggg	cca	2261
Gln	Pro	Arg	Asn	Ile	Ala	Lys	Arg	His	Asp	Glu	Gly	His	Lys	Gly	Pro	
	570					575					580					
aga	gat	ctc	cag	gct	tcc	agt	ggt	ggc	aac	agg	ggc	agg	atg	ctg	gca	2309
Arg	Asp	Leu	Gln	Ala	Ser	Ser	Gly	Gly	Asn	Arg	Gly	Arg	Met	Leu	Ala	
585					590				595					600		
gat	agc	agc	aac	gcc	gtg	ggc	cca	cct	acc	act	gtc	cga	gtg	aca	cac	2357
Asp	Ser	Ser	Asn	Ala	Val	Gly	Pro	Pro	Thr	Thr	Val	Arg	Val	Thr	His	
				605					610					615		
aag	tgt	ttt	att	ctt	ccc	aat	gac	tct	atc	cat	tgt	gag	aga	gaa	ctg	2405
Lys	Cys	Phe	Ile	Leu	Pro	Asn	Asp	Ser	Ile	His	Cys	Glu	Arg	Glu	Leu	
			620					625					630			
tac	caa	tcg	gcc	aga	gcg	tgg	aag	gac	cat	aag	gca	tac	att	gac	aaa	2453
Tyr	Gln	Ser	Ala	Arg	Ala	Trp	Lys	Asp	His	Lys	Ala	Tyr	Ile	Asp	Lys	
		635					640					645				
gag	att	gaa	gct	ctg	caa	gat	aaa	att	aag	aat	tta	aga	gaa	gtg	aga	2501
Glu	Ile	Glu	Ala	Leu	Gln	Asp	Lys	Ile	Lys	Asn	Leu	Arg	Glu	Val	Arg	
	650					655					660					
gga	cat	ctg	aag	aga	agg	aag	cct	gag	gaa	tgt	agc	tgc	agt	aaa	caa	2549
Gly	His	Leu	Lys	Arg	Arg	Lys	Pro	Glu	Glu	Cys	Ser	Cys	Ser	Lys	Gln	
665				670						675					680	
agc	tat	tac	aat	aaa	gag	aaa	ggt	gta	aaa	aag	caa	gag	aaa	tta	aag	2597
Ser	Tyr	Tyr	Asn	Lys	Glu	Lys	Gly	Val	Lys	Lys	Gln	Glu	Lys	Leu	Lys	
				685				690						695		
agc	cat	ctt	cac	cca	ttc	aag	gag	gct	gct	cag	gaa	gta	gat	agc	aaa	2645
Ser	His	Leu	His	Pro	Phe	Lys	Glu	Ala	Ala	Gln	Glu	Val	Asp	Ser	Lys	
			700					705					710			

## 5800-79B SEQLIST.TXT

ctg caa ctt ttc aag gag aac aac cgt agg agg aag aag gag agg aag	2693
Leu Gln Leu Phe Lys Glu Asn Asn Arg Arg Arg Lys Lys Glu Arg Lys	
715 720 725	
gag aag aga cgg cag agg aag ggg gaa gag tgc agc ctg cct ggc ctc	2741
Glu Lys Arg Arg Gln Arg Lys Gly Glu Glu Cys Ser Leu Pro Gly Leu	
730 735 740	
act tgc ttc acg cat gac aac aac cac tgg cag aca gcc ccg ttc tgg	2789
Thr Cys Phe Thr His Asp Asn Asn His Trp Gln Thr Ala Pro Phe Trp	
745 750 755 760	
aac ctg gga tct ttc tgt gct tgc acg agt tct aac aat aac acc tac	2837
Asn Leu Gly Ser Phe Cys Ala Cys Thr Ser Ser Asn Asn Asn Thr Tyr	
765 770 775	
tgg tgt ttg cgt aca gtt aat gag acg cat aat ttt ctt ttc tgt gag	2885
Trp Cys Leu Arg Thr Val Asn Glu Thr His Asn Phe Leu Phe Cys Glu	
780 785 790	
ttt gct act ggc ttt ttg gag tat ttt gat atg aat aca gat cct tat	2933
Phe Ala Thr Gly Phe Leu Glu Tyr Phe Asp Met Asn Thr Asp Pro Tyr	
795 800 805	
cag ctc aca aat aca gtg cac acg gta gaa cga ggc att ttg aat cag	2981
Gln Leu Thr Asn Thr Val His Thr Val Glu Arg Gly Ile Leu Asn Gln	
810 815 820	
cta cac gta caa cta atg gag ctc aga agc tgt caa gga tat aag cag	3029
Leu His Val Gln Leu Met Glu Leu Arg Ser Cys Gln Gly Tyr Lys Gln	
825 830 835 840	
tgc aac cca aga cct aag aat ctt gat gtt gga aat aaa gat gga gga	3077
Cys Asn Pro Arg Pro Lys Asn Leu Asp Val Gly Asn Lys Asp Gly Gly	
845 850 855	
agc tat gac cta cac aga gga cag tta tgg gat gga tgg gaa ggt taa	3125
Ser Tyr Asp Leu His Arg Gly Gln Leu Trp Asp Gly Trp Glu Gly *	
860 865 870	
tcagccccgt ctactgacg acatcaactg gcaaggccta gaggagctac acagtgtgaa	3185
tgaaaacatc tatgagtaca gacaaaacta cagacttagt ctggtggact ggactaatta	3245
cttgaaggat ttagatagag tatttgcact gctgaagagt cactatgagc aaaataaaac	3305
aaataagact caaactgctc aaagtgcagg gttcttggtt gtctctgctg agcacgctgt	3365
gtcaatggag atggcctctg ctgactcaga tgaagaccca aggcataagg ttgggaaaac	3425
acctcatttg accttgccag ctgaccttca aaccctgcat ttgaaccgac caacattaag	3485
tccagagagt aaacttgaat ggaataacga cattccagaa gttaatcatt tgaattctga	3545
acactggaga aaaaccgaaa aatggacggg gcatgaagag actaatcatc tggaaaccga	3605
tttcagtggc gatggcatga cagagctaga gctcgggccc agccccaggc tgcagcccac	3665
tcgcaggcac ccgaaagaac ttccccagta tgggtggcct ggaaaggaca tttttgaaga	3725
tcaactatat cticctgtgc attccgatgg aatttcagtt catcagatgt tcaccatggc	3785
caccgcagaa caccgaagta attccagcat agcggggaag atgttgacca aggtggagaa	3845
gaatcacgaa aaggagaagt cacagcacct agaaggcagc gcctcctctt cactctcctc	3905
tgattagatg aaactgttac cttaccctaa acacagtatt tctttttaac ttttttattt	3965
gtaaactaat aaaggkaatc acagccacca acattccaag ctaccctggg tacctttgtg	4025
cagtagaagc tagtgagcat gtgagcaagc ggtgtgcaca cggagactca tcgttataat	4085
ttactactctg ccaaggagta gaaagaaagg ctggggatat ttgggttggc tttggktttg	4145
attttttgct tgggtgggtg gtttgkacta aaacagtatt atcttttgaa tatcgtaggg	4205
acataarkww wwwmmwkkw wtcmawymra kakgsywrri wkgggstyty tsckrkstmw	4265
amwykwscmc cyskkrrwwaw tywywmmywc mykytssstg rykrnktaat gaagtt	4321

## 5800-79B SEQLIST.TXT

<211> 569  
 <212> PRT  
 <213> homo sapiens

<400> 5  
 Met His Thr Leu Thr Gly Phe Ser Leu Val Ser Leu Leu Ser Phe Gly  
 1 5 10 15  
 Tyr Leu Ser Trp Asp Trp Ala Lys Pro Ser Phe Val Ala Asp Gly Pro  
 20 25 30  
 Gly Glu Ala Gly Glu Gln Pro Ser Ala Ala Pro Pro Gln Pro Pro His  
 35 40 45  
 Ile Ile Phe Ile Leu Thr Asp Asp Gln Gly Tyr His Asp Val Gly Tyr  
 50 55 60  
 His Gly Ser Asp Ile Glu Thr Pro Thr Leu Asp Arg Leu Ala Ala Lys  
 65 70 75 80  
 Gly Val Lys Leu Glu Asn Tyr Tyr Ile Gln Pro Ile Cys Thr Pro Ser  
 85 90 95  
 Arg Ser Gln Leu Leu Thr Gly Arg Tyr Gln Ile His Thr Gly Leu Gln  
 100 105 110  
 His Ser Ile Ile Arg Pro Gln Gln Pro Asn Cys Leu Pro Leu Asp Gln  
 115 120 125  
 Val Thr Leu Pro Gln Lys Leu Gln Glu Ala Gly Tyr Ser Thr His Met  
 130 135 140  
 Val Gly Lys Trp His Leu Gly Phe Tyr Arg Lys Glu Cys Leu Pro Thr  
 145 150 155 160  
 Arg Arg Gly Phe Asp Thr Phe Leu Gly Ser Leu Thr Gly Asn Val Asp  
 165 170 175  
 Tyr Tyr Thr Tyr Asp Asn Cys Asp Gly Pro Gly Val Cys Gly Phe Asp  
 180 185 190  
 Leu His Glu Gly Glu Asn Val Ala Trp Gly Leu Ser Gly Gln Tyr Ser  
 195 200 205  
 Thr Met Leu Tyr Ala Gln Arg Ala Ser His Ile Leu Ala Ser His Ser  
 210 215 220  
 Pro Gln Arg Pro Leu Phe Leu Tyr Val Ala Phe Gln Ala Val His Thr  
 225 230 235 240  
 Pro Leu Gln Ser Pro Arg Glu Tyr Leu Tyr Arg Tyr Arg Thr Met Gly  
 245 250 255  
 Asn Val Ala Arg Arg Lys Tyr Ala Ala Met Val Thr Cys Met Asp Glu  
 260 265 270  
 Ala Val Arg Asn Ile Thr Trp Ala Leu Lys Arg Tyr Gly Phe Tyr Asn  
 275 280 285  
 Asn Ser Val Ile Ile Phe Ser Ser Asp Asn Gly Gly Gln Thr Phe Ser  
 290 295 300  
 Gly Gly Ser Asn Trp Pro Leu Arg Gly Arg Lys Gly Thr Tyr Trp Glu  
 305 310 315 320  
 Gly Gly Val Arg Gly Leu Gly Phe Val His Ser Pro Leu Leu Lys Arg  
 325 330 335  
 Lys Gln Arg Thr Ser Arg Ala Leu Met His Ile Thr Asp Trp Tyr Pro  
 340 345 350  
 Thr Leu Val Gly Leu Ala Gly Gly Thr Thr Ser Ala Ala Asp Gly Leu  
 355 360 365  
 Asp Gly Tyr Asp Val Trp Pro Ala Ile Ser Glu Gly Arg Ala Ser Pro  
 370 375 380  
 Arg Thr Glu Ile Leu His Asn Ile Asp Pro Leu Tyr Asn His Ala Gln  
 385 390 395 400  
 His Gly Ser Leu Glu Gly Gly Phe Gly Ile Trp Asn Thr Ala Val Gln  
 405 410 415  
 Ala Ala Ile Arg Val Gly Glu Trp Lys Leu Leu Thr Gly Asp Pro Gly  
 420 425 430  
 Tyr Gly Asp Trp Ile Pro Pro Gln Thr Leu Ala Thr Phe Pro Gly Ser  
 435 440 445  
 Trp Trp Asn Leu Glu Arg Met Ala Ser Val Arg Gln Ala Val Trp Leu  
 450 455 460

5800-79B SEQLIST.TXT

Phe	Asn	Ile	Ser	Ala	Asp	Pro	Tyr	Glu	Arg	Glu	Asp	Leu	Ala	Gly	Gln
465					470					475					480
Arg	Pro	Asp	Val	Val	Arg	Thr	Leu	Leu	Ala	Arg	Leu	Ala	Glu	Tyr	Asn
				485					490					495	
Arg	Thr	Ala	Ile	Pro	Val	Arg	Tyr	Pro	Ala	Glu	Asn	Pro	Arg	Ala	His
			500					505					510		
Pro	Asp	Phe	Asn	Gly	Gly	Ala	Trp	Gly	Pro	Trp	Ala	Ser	Asp	Glu	Glu
		515					520					525			
Glu	Glu	Glu	Glu	Glu	Gly	Arg	Ala	Arg	Ser	Phe	Ser	Arg	Gly	Arg	Arg
	530					535					540				
Lys	Lys	Lys	Cys	Lys	Ile	Cys	Lys	Leu	Arg	Ser	Phe	Phe	Arg	Lys	Leu
545					550					555					560
Asn	Thr	Arg	Leu	Met	Ser	Gln	Arg	Ile							
				565											

<210> 6  
 <211> 2940  
 <212> DNA  
 <213> homo sapiens

<220>  
 <221> CDS  
 <222> (334)...(2043)

<400> 6

ccacgcgtcc	gccccacgcgt	ccggctgcca	cgccgcgtct	caggctggcc	gggctgagcc		60
ggggaagagg	gagcaaaggc	ggcgcagggc	ctgcgcttag	gcagcgggag	gcagctcggc		120
gcgggcctga	cctccccaga	gcgccccgct	gcggccgagc	agatccggcc	cagccgtccg		180
gcagccagtc	ccggaccaga	cactggaccg	tccccggggg	gcgctgaact	ccctcgcagc		240
atccgagccg	gcgggcccgt	ggtgcgccct	gggcgcgcga	ggtggtgagg	ccccaggagc		300
ccggcgcgcc	gggacacgcg	ggccggcttg	gcg atg cac	acc ctc act	ggc ttc		354
			Met His	Thr Leu	Thr Gly Phe		
			1		5		
tct ctg gtc	agc ctg ctc	agc ttc	ggc tac	ctg tcc	tgg gac	tgg gcc	402
Ser Leu Val	Ser Leu Leu	Ser Phe	Gly Tyr	Leu Ser	Trp Asp	Trp Ala	
	10	15		20			
aag ccg agc	ttc gtg gcc	gac ggg	ccc ggg	gag gct	ggc gag	cag ccc	450
Lys Pro Ser	Phe Val Ala	Asp Gly	Pro Gly	Glu Ala	Gly Glu	Gln Pro	
	25	30		35			
tcg gcc gct	ccg ccc cag	cct ccc	cac atc	atc atc	ttc atc	ctc acg	498
Ser Ala Ala	Pro Pro Gln	Pro Pro	His Ile	Ile Ile	Phe Ile	Leu Thr	
	40	45		50		55	
gac caa ggc	tac cac gac	gtg ggc	tac cat	ggt tca	gat atc	gag acc	546
Asp Gln Gly	Tyr His Asp	Val Gly	Tyr His	Gly Ser	Asp Ile	Glu Thr	
	60	65		70			
cct acg ctg	gac agg ctg	gcg gcc	aag ggg	gtc aag	ttg gag	aat tat	594
Pro Thr Leu	Asp Arg Leu	Ala Ala	Lys Gly	Val Lys	Leu Glu	Asn Tyr	
	75	80		85			
tac atc cag	ccc atc tgc	acg cct	tcg cgg	agc cag	ctc ctc	act ggc	642
Tyr Ile Gln	Pro Ile Cys	Thr Pro	Ser Arg	Ser Gln	Leu Leu	Thr Gly	
	90	95		100			
agg tac cag	atc cac aca	gga ctc	cag cat	tcc atc	atc cgc	cca cag	690
Arg Tyr Gln	Ile His Thr	Gly Leu	Gln His	Ser Ile	Ile Arg	Pro Gln	
	105	110		115			

## 5800-79B SEQLIST.TXT

cag Gln 120	ccc Pro	aac Asn	tgc Cys	ctg Leu	ccc Pro 125	ctg Leu	gac Asp	cag Gln	gtg Val 130	aca Thr	ctg Leu	cca Pro	cag Gln	aag Lys	ctg Leu 135	738
cag Gln	gag Glu	gca Ala	ggt Gly 140	tat Tyr 140	tcc Ser	acc Thr	cat His	atg Met	gtg Val 145	ggc Gly	aag Lys	tgg Trp	cac His	ctg Leu 150	ggc Gly	786
ttc Phe	tac Tyr	cgg Arg	aag Lys 155	gag Glu	tgt Cys	ctg Leu	ccc Pro	acc Thr 160	cgt Arg	cgg Arg	ggc Gly	ttc Phe	gac Asp 165	acc Thr	ttc Phe	834
ctg Leu	ggc Gly	tgc Ser 170	ctc Leu	acg Thr	ggc Gly	aat Asn	gtg Val 175	gac Asp	tat Tyr	tac Tyr	acc Thr	tat Tyr 180	gac Asp	aac Asn	tgt Cys	882
gat Asp 185	ggc Gly	cca Pro	ggc Gly	gtg Val	tgc Cys 190	ggc Gly	ttc Phe	gac Asp	ctg Leu	cac His	gag Glu 195	ggt Gly	gag Glu	aat Asn	gtg Val	930
gcc Ala 200	tgg Trp	ggg Gly	ctc Leu	agc Ser	ggc Gly 205	cag Gln	tac Tyr	tcc Ser	act Thr	atg Met 210	ctt Leu	tac Tyr	gcc Ala	cag Gln	cgc Arg 215	978
gcc Ala	agc Ser	cat His	atc Ile	ctg Leu 220	gcc Ala	agc Ser	cac His	agc Ser	cct Pro 225	cag Gln	cgt Arg	ccc Pro	ctc Leu	ttc Phe 230	ctc Leu	1026
tat Tyr	gtg Val	gcc Ala	ttc Phe 235	cag Gln	gca Ala	gta Val	cac His	aca Thr 240	ccc Pro	ctg Leu	cag Gln	tcc Ser	cct Pro 245	cgt Arg	gag Glu	1074
tac Tyr	ctg Leu	tac Tyr 250	cgc Arg	tac Tyr	cgc Arg	acc Thr	atg Met 255	ggc Gly	aat Asn	gtg Val	gcc Ala	cgg Arg 260	cgg Arg	aag Lys	tac Tyr	1122
gcg Ala 265	gcc Ala	atg Met	gtg Val	acc Thr	tgc Cys	atg Met 270	gat Asp	gag Glu	gct Ala	gtg Val	cgc Arg 275	aac Asn	atc Ile	acc Thr	tgg Trp	1170
gcc Ala 280	ctc Leu	aag Lys	cgc Arg	tac Tyr	ggt Gly 285	ttc Phe	tac Tyr	aac Asn	aac Asn	agt Ser 290	gtc Val	atc Ile	atc Ile	ttc Phe	tcc Ser 295	1218
agt Ser	gac Asp	aat Asn	ggt Gly 300	ggc Gly	cag Gln	act Thr	ttc Phe	tgc Ser	ggg Gly 305	ggc Gly	agc Ser	aac Asn	tgg Trp	ccg Pro 310	ctc Leu	1266
cga Arg	gga Gly	cgc Arg	aag Lys 315	ggc Gly	act Thr	tat Tyr	tgg Trp	gaa Glu 320	ggt Gly	ggc Gly	gtg Val	cgg Arg	ggc Gly 325	cta Leu	ggc Gly	1314
ttt Phe	gtc Val	cac His 330	agt Ser	ccc Pro	ctg Leu	ctc Leu	aag Lys 335	cga Arg	aag Lys	caa Gln	cgg Arg	aca Thr 340	agc Ser	cgg Arg	gca Ala	1362
ctg Leu	atg Met 345	cac His	atc Ile	act Thr	gac Asp	tgg Trp 350	tac Tyr	ccg Pro	acc Thr	ctg Leu	gtg Val 355	ggt Gly	ctg Leu	gca Ala	ggt Gly	1410
ggt Gly 360	acc Thr	acc Thr	tca Ser	gca Ala	gcc Ala 365	gat Asp	ggg Gly	cta Leu	gat Asp	ggc Gly 370	tac Tyr	gac Asp	gtg Val	tgg Trp	ccg Pro 375	1458

## 5800-79B SEQLIST.TXT

gcc atc agc gag ggc cgg gcc tca cca cgc acg gag atc ctg cac aac	1506
Ala Ile Ser Glu Gly Arg Ala Ser Pro Arg Thr Glu Ile Leu His Asn	
380 385 390	
att gac cca ctc tac aac cat gcc cag cat ggc tcc ctg gag ggc ggc	1554
Ile Asp Pro Leu Tyr Asn His Ala Gln His Gly Ser Leu Glu Gly Gly	
395 400 405	
ttt ggc atc tgg aac acc gcc gtg cag gct gcc atc cgc gtg ggt gag	1602
Phe Gly Ile Trp Asn Thr Ala Val Gln Ala Ala Ile Arg Val Gly Glu	
410 415 420	
tgg aag ctg ctg aca gga gac ccc ggc tat ggc gat tgg atc cca ccg	1650
Trp Lys Leu Leu Thr Gly Asp Pro Gly Tyr Gly Asp Trp Ile Pro Pro	
425 430 435	
cag aca ctg gcc acc ttc ccg ggt agc tgg tgg aac ctg gaa cga atg	1698
Gln Thr Leu Ala Thr Phe Pro Gly Ser Trp Trp Asn Leu Glu Arg Met	
440 445 450 455	
gcc agt gtc cgc cag gcc gtg tgg ctc ttc aac atc agt gct gac cct	1746
Ala Ser Val Arg Gln Ala Val Trp Leu Phe Asn Ile Ser Ala Asp Pro	
460 465 470	
tat gaa cgg gag gac ctg gct ggc cag cgg cct gat gtg gtc cgc acc	1794
Tyr Glu Arg Glu Asp Leu Ala Gly Gln Arg Pro Asp Val Val Arg Thr	
475 480 485	
ctg ctg gct cgc ctg gcc gaa tat aac cgc aca gcc atc ccg gta cgc	1842
Leu Leu Ala Arg Leu Ala Glu Tyr Asn Arg Thr Ala Ile Pro Val Arg	
490 495 500	
tac cca gct gag aac ccc cgg gct cat cct gac ttt aat ggg ggt gct	1890
Tyr Pro Ala Glu Asn Pro Arg Ala His Pro Asp Phe Asn Gly Gly Ala	
505 510 515	
tgg ggg ccc tgg gcc agt gat gag gaa gag gag gaa gag gaa ggg agg	1938
Trp Gly Pro Trp Ala Ser Asp Glu Glu Glu Glu Glu Glu Glu Gly Arg	
520 525 530 535	
gct cga agc ttc tcc cgg ggt cgt cgc aag aaa aaa tgc aag att tgc	1986
Ala Arg Ser Phe Ser Arg Gly Arg Arg Lys Lys Lys Cys Lys Ile Cys	
540 545 550	
aag ctt cga tcc ttt ttc cgt aaa ctc aac acc agg cta atg tcc caa	2034
Lys Leu Arg Ser Phe Phe Arg Lys Leu Asn Thr Arg Leu Met Ser Gln	
555 560 565	
cgg atc tga tgggtggggag ggagaaaact gtcctttaga ggatcttccc	2083
Arg Ile *	

cactccggct tggccctgct gtttctcagg gagaagcctg tcacatctcc atctacaggg	2143
agttggaggg tgtagagtc cttggttgaa cagggtaggg agcctggata ggagtgggtg	2203
ggaataaacc agactgggat gcctgtgtct cagtcctgcc tcctcacgga cttgctctgt	2263
gacctcaggt gacctacatg agcttttagc ctcagtticc tcactctgtaa aatgagctct	2323
aatgactttg tgactctttg gtgtggccct ggagcctggg gccacgggtg agttcctggc	2383
cggccttgcc acttgacaac tcctttaagg cttccccctt aacacgggat ccctgtggtg	2443
gtgtttggga gttgcctgga ggcaactcca agcctggccc ccagctgaag catggcaatc	2503
tggctgctct ctacagggac ccccaagcgc tgtgggtgga gggcaggggt cggggggggt	2563
gaccttcttg ggtcttcaca tggcctaggc cagtcctccg gtcagactgg tgtcaggcac	2623
cgtggtgcaa aattcctctt ctggccctc cagtaccag agaaaactggc tgggccatta	2683

## 5800-79B SEQLIST.TXT

actgctgcag	caccaagggg	ggtagaaaga	gctgtgaaga	gccccaaac	cagtaccagg	2743
acacctgggt	tctcctgtga	cctggggcac	agttcttgcc	ctctaggcct	tgatttcccc	2803
acctgcaagt	ggggatgcc	gccctggctc	tgcctccttc	atgaggctct	ggaagactgg	2863
ccaaggttgt	ggaggagctt	gtgaacttga	ttaaagtgtc	gtaacatgga	aaaaaaaaa	2923
aaaaaaaaa	agggcg					2940

<210> 7  
 <211> 599  
 <212> PRT  
 <213> homo sapiens

<400> 7

Met	Ala	Pro	Arg	Gly	Cys	Ala	Gly	His	Pro	Pro	Pro	Pro	Ser	Pro	Gln
1				5					10					15	
Ala	Cys	Val	Cys	Pro	Gly	Lys	Met	Leu	Ala	Met	Gly	Ala	Leu	Ala	Gly
			20					25					30		
Phe	Trp	Ile	Leu	Cys	Leu	Leu	Thr	Tyr	Gly	Tyr	Leu	Ser	Trp	Gly	Gln
		35					40					45			
Ala	Leu	Glu	Glu	Glu	Glu	Gly	Ala	Leu	Leu	Ala	Gln	Ala	Gly	Glu	
	50				55					60					
Lys	Leu	Glu	Pro	Ser	Thr	Thr	Ser	Thr	Ser	Gln	Pro	His	Leu	Ile	Phe
65					70				75					80	
Ile	Leu	Ala	Asp	Asp	Gln	Gly	Phe	Arg	Asp	Val	Gly	Tyr	His	Gly	Ser
			85					90					95		
Glu	Ile	Lys	Thr	Pro	Thr	Leu	Asp	Lys	Leu	Ala	Ala	Glu	Gly	Val	Lys
			100					105					110		
Leu	Glu	Asn	Tyr	Tyr	Val	Gln	Pro	Ile	Cys	Thr	Pro	Ser	Arg	Ser	Gln
		115					120					125			
Phe	Ile	Thr	Gly	Lys	Tyr	Gln	Ile	His	Thr	Gly	Leu	Gln	His	Ser	Ile
	130					135					140				
Ile	Arg	Pro	Thr	Gln	Pro	Asn	Cys	Leu	Pro	Leu	Asp	Asn	Ala	Thr	Leu
145					150					155					160
Pro	Gln	Lys	Leu	Lys	Glu	Val	Gly	Tyr	Ser	Thr	His	Met	Val	Gly	Lys
				165					170					175	
Trp	His	Leu	Gly	Phe	Tyr	Arg	Lys	Glu	Cys	Met	Pro	Thr	Arg	Arg	Gly
			180					185					190		
Phe	Asp	Thr	Phe	Phe	Gly	Ser	Leu	Leu	Gly	Ser	Gly	Asp	Tyr	Tyr	Thr
		195					200					205			
His	Tyr	Lys	Cys	Asp	Ser	Pro	Gly	Met	Cys	Gly	Tyr	Asp	Leu	Tyr	Glu
	210					215					220				
Asn	Asp	Asn	Ala	Ala	Trp	Asp	Tyr	Asp	Asn	Gly	Ile	Tyr	Ser	Thr	Gln
225					230					235					240
Met	Tyr	Thr	Gln	Arg	Val	Gln	Gln	Ile	Leu	Ala	Ser	His	Asn	Pro	Thr
				245					250					255	
Lys	Pro	Ile	Phe	Leu	Tyr	Ile	Ala	Tyr	Gln	Ala	Val	His	Ser	Pro	Leu
			260					265					270		
Gln	Ala	Pro	Gly	Arg	Tyr	Phe	Glu	His	Tyr	Arg	Ser	Ile	Ile	Asn	Ile
		275					280					285			
Asn	Arg	Arg	Arg	Tyr	Ala	Ala	Met	Leu	Ser	Cys	Leu	Asp	Glu	Ala	Ile
	290					295					300				
Asn	Asn	Val	Thr	Leu	Ala	Leu	Lys	Thr	Tyr	Gly	Phe	Tyr	Asn	Asn	Ser
305					310					315					320
Ile	Ile	Ile	Tyr	Ser	Ser	Asp	Asn	Gly	Gly	Gln	Pro	Thr	Ala	Gly	Gly
				325					330					335	
Ser	Asn	Trp	Pro	Leu	Arg	Gly	Ser	Lys	Gly	Thr	Tyr	Trp	Glu	Gly	Gly
			340					345					350		
Ile	Arg	Ala	Val	Gly	Phe	Val	His	Ser	Pro	Leu	Leu	Lys	Asn	Lys	Gly
		355					360					365			
Thr	Val	Cys	Lys	Glu	Leu	Val	His	Ile	Thr	Asp	Trp	Tyr	Pro	Thr	Leu
	370					375					380				
Ile	Ser	Leu	Ala	Glu	Gly	Gln	Ile	Asp	Glu	Asp	Ile	Gln	Leu	Asp	Gly
385					390					395					400
Tyr	Asp	Ile	Trp	Glu	Thr	Ile	Ser	Glu	Gly	Leu	Arg	Ser	Pro	Arg	Val

## 5800-79B SEQLIST.TXT

405 410 415  
 Asp Ile Leu His Asn Ile Asp Pro Ile Tyr Thr Lys Ala Lys Asn Gly  
 420 425 430  
 Ser Trp Ala Ala Gly Tyr Gly Ile Trp Asn Thr Ala Ile Gln Ser Ala  
 435 440 445  
 Ile Arg Val Gln His Trp Lys Leu Leu Thr Gly Asn Pro Gly Tyr Ser  
 450 455 460  
 Asp Trp Val Pro Pro Gln Ser Phe Ser Asn Leu Gly Pro Asn Arg Trp  
 465 470 475 480  
 His Asn Glu Arg Ile Thr Leu Ser Thr Gly Lys Ser Val Trp Leu Phe  
 485 490 495  
 Asn Ile Thr Ala Asp Pro Tyr Glu Arg Val Asp Leu Ser Asn Arg Tyr  
 500 505 510  
 Pro Gly Ile Val Lys Lys Leu Leu Arg Arg Leu Ser Gln Phe Asn Lys  
 515 520 525  
 Thr Ala Val Pro Val Arg Tyr Pro Pro Lys Asp Pro Arg Ser Asn Pro  
 530 535 540  
 Arg Leu Asn Gly Gly Val Trp Gly Pro Trp Tyr Lys Glu Glu Thr Lys  
 545 550 555 560  
 Lys Lys Lys Pro Ser Lys Asn Gln Ala Glu Lys Lys Gln Lys Lys Ser  
 565 570 575  
 Lys Lys Lys Lys Lys Lys Gln Gln Lys Ala Val Ser Gly Ser Thr Cys  
 580 585 590  
 His Ser Gly Val Thr Cys Gly  
 595

<210> 8  
 <211> 2253  
 <212> DNA  
 <213> homo sapiens

<220>  
 <221> CDS  
 <222> (324)...(2123)

<400> 8  
 cacgctccg cccacgcgtc cgtggagata ttaacttttt tctttttttt tttccttggt 60  
 ggaagctgct ctaggaggagg gggaggagga ggagaaagt aaatgtgctg gagaagagcg 120  
 agccctcctt gttcttccgg agtcccattc attagccat cacttctgga agattaaagt 180  
 tgtcggacat ggtgacagct gagaggagag gaggatttct tgccagggtg agagtcttca 240  
 ccgtctgttg ggtgcatgtg tgcgcccgca gcggcgcggg gcgctggtt ctccgcgtgg 300  
 agtctcacct gggacctgag tga atg gct ccc agg ggc tgt gcg ggg cat ccg 353  
 Met Ala Pro Arg Gly Cys Ala Gly His Pro  
 1 5 10  
 cct ccg cct tct cca cag gcc tgt gtc tgt cct gga aag atg cta gca 401  
 Pro Pro Pro Ser Pro Gln Ala Cys Val Cys Pro Gly Lys Met Leu Ala  
 15 20 25  
 atg ggg gcg ctg gca gga ttc tgg atc ctc tgc ctc ctc act tat ggt 449  
 Met Gly Ala Leu Ala Gly Phe Trp Ile Leu Cys Leu Leu Thr Tyr Gly  
 30 35 40  
 tac ctg tcc tgg ggc cag gcc tta gaa gag gag gaa gaa ggg gcc tta 497  
 Tyr Leu Ser Trp Gly Gln Ala Leu Glu Glu Glu Glu Gly Ala Leu  
 45 50 55  
 cta gct caa gct gga gag aaa cta gag ccc agc aca act tcc acc tcc 545  
 Leu Ala Gln Ala Gly Glu Lys Leu Glu Pro Ser Thr Thr Ser Thr Ser  
 60 65 70  
 cag ccc cat ctc att ttc atc cta gcg gat gat cag gga ttt aga gat 593  
 Page 16



5800-79B SEQLIST.TXT

Gln 75	Pro	His	Leu	Ile	Phe 80	Ile	Leu	Ala	Asp	Asp 85	Gln	Gly	Phe	Arg	Asp 90	
gtg Val	ggt Gly	tac Tyr	cac His	gga Gly 95	tct Ser	gag Glu	att Ile	aaa Lys	aca Thr 100	cct Pro	act Thr	ctt Leu	gac Asp	aag Lys 105	ctc Leu	641
gct Ala	gcc Ala	gaa Glu	gga Gly 110	gtt Val	aaa Lys	ctg Leu	gag Glu	aac Asn 115	tac Tyr	tat Tyr	gtc Val	cag Gln	cct Pro 120	att Ile	tgc Cys	689
aca Thr	cca Pro	tcc Ser 125	agg Arg	agt Ser	cag Gln	ttt Phe	att Ile 130	act Thr	gga Gly	aag Lys	tat Tyr	cag Gln 135	ata Ile	cac His	acc Thr	737
gga Gly	ctt Leu 140	caa Gln	cat His	tct Ser	atc Ile	ata Ile 145	aga Arg	cct Pro	acc Thr	caa Gln	ccc Pro 150	aac Asn	tgt Cys	tta Leu	cct Pro	785
ctg Leu 155	gac Asp	aat Asn	gcc Ala	acc Thr	cta Leu 160	cct Pro	cag Gln	aaa Lys	ctg Leu	aag Lys 165	gag Glu	gtt Val	gga Gly	tat Tyr	tca Ser 170	833
acg Thr	cat His	atg Met	gtc Val	gga Gly 175	aaa Lys	tgg Trp	cac His	ttg Leu	ggt Gly 180	ttt Phe	tac Tyr	aga Arg	aaa Lys	gaa Glu 185	tgc Cys	881
atg Met	ccc Pro	acc Thr	aga Arg 190	aga Arg	gga Gly	ttt Phe	gat Asp	acc Thr 195	ttt Phe	ttt Phe	ggt Gly	tcc Ser	ctt Leu 200	ttg Leu	gga Gly	929
agt Ser	ggg Gly	gat Asp 205	tac Tyr	tat Tyr	aca Thr	cac His	tac Tyr 210	aaa Lys	tgt Cys	gac Asp	agt Ser	cct Pro 215	ggg Gly	atg Met	tgt Cys	977
ggc Gly	tat Tyr 220	gac Asp	ttg Leu	tat Tyr	gaa Glu	aac Asn 225	gac Asp	aat Asn	gct Ala	gcc Ala	tgg Trp 230	gac Asp	tat Tyr	gac Asp	aat Asn	1025
ggc Gly 235	ata Ile	tac Tyr	tcc Ser	aca Thr	cag Gln 240	atg Met	tac Tyr	act Thr	cag Gln	aga Arg 245	gta Val	cag Gln	caa Gln	atc Ile	tta Leu 250	1073
gct Ala	tcc Ser	cat His	aac Asn	ccc Pro 255	aca Thr	aag Lys	cct Pro	ata Ile	ttt Phe 260	tta Leu	tat Tyr	att Ile	gcc Ala	tat Tyr 265	caa Gln	1121
gct Ala	gtt Val	cat His	tca Ser 270	cca Pro	ctg Leu	caa Gln	gct Ala	cct Pro 275	ggc Gly	agg Arg	tat Tyr	ttc Phe	gaa Glu 280	cac His	tac Tyr	1169
cga Arg	tcc Ser	att Ile 285	atc Ile	aac Asn	ata Ile	aac Asn	agg Arg 290	agg Arg	aga Arg	tat Tyr	gct Ala	gcc Ala 295	atg Met	ctt Leu	tcc Ser	1217
tgc Cys	tta Leu 300	gat Asp	gaa Glu	gca Ala	atc Ile	aac Asn 305	aac Asn	gtg Val	aca Thr	ttg Leu	gct Ala 310	cta Leu	aag Lys	act Thr	tat Tyr	1265
ggt Gly 315	ttc Phe	tat Tyr	aac Asn	aac Asn	agc Ser 320	att Ile	atc Ile	att Ile	tac Tyr	tct Ser 325	tca Ser	gat Asp	aat Asn	ggt Gly	ggc Gly 330	1313

5800-79B SEQLIST.TXT

cag	cct	acg	gca	gga	ggg	agt	aac	tgg	cct	ctc	aga	ggt	agc	aaa	gga	1361
Gln	Pro	Thr	Ala	Gly	Gly	Ser	Asn	Trp	Pro	Leu	Arg	Gly	Ser	Lys	Gly	
				335					340					345		
aca	tat	tgg	gaa	gga	ggg	atc	cgg	gct	gta	ggc	ttt	gtg	cat	agc	cca	1409
Thr	Tyr	Trp	Glu	Gly	Gly	Ile	Arg	Ala	Val	Gly	Phe	Val	His	Ser	Pro	
			350					355					360			
ctt	ctg	aaa	aac	aag	gga	aca	gtg	tgt	aag	gaa	ctt	gtg	cac	atc	act	1457
Leu	Leu	Lys	Asn	Lys	Gly	Thr	Val	Cys	Lys	Glu	Leu	Val	His	Ile	Thr	
		365					370					375				
gac	tgg	tac	ccc	act	ctc	att	tca	ctg	gct	gaa	gga	cag	att	gat	gag	1505
Asp	Trp	Tyr	Pro	Thr	Leu	Ile	Ser	Leu	Ala	Glu	Gly	Gln	Ile	Asp	Glu	
	380					385					390					
gac	att	caa	cta	gat	ggc	tat	gat	atc	tgg	gag	acc	ata	agt	gag	ggt	1553
Asp	Ile	Gln	Leu	Asp	Gly	Tyr	Asp	Ile	Trp	Glu	Thr	Ile	Ser	Glu	Gly	
	395				400					405					410	
ctt	cgc	tca	ccc	cga	gta	gat	att	ttg	cat	aac	att	gac	ccc	ata	tac	1601
Leu	Arg	Ser	Pro	Arg	Val	Asp	Ile	Leu	His	Asn	Ile	Asp	Pro	Ile	Tyr	
				415					420					425		
acc	aag	gca	aaa	aat	ggc	tcc	tgg	gca	gca	ggc	tat	ggg	atc	tgg	aac	1649
Thr	Lys	Ala	Lys	Asn	Gly	Ser	Trp	Ala	Ala	Gly	Tyr	Gly	Ile	Trp	Asn	
			430					435					440			
act	gca	atc	cag	tca	gcc	atc	aga	gtg	cag	cac	tgg	aaa	ttg	ctt	aca	1697
Thr	Ala	Ile	Gln	Ser	Ala	Ile	Arg	Val	Gln	His	Trp	Lys	Leu	Leu	Thr	
		445					450					455				
gga	aat	cct	ggc	tac	agc	gac	tgg	gtc	ccc	cct	cag	tct	ttc	agc	aac	1745
Gly	Asn	Pro	Gly	Tyr	Ser	Asp	Trp	Val	Pro	Pro	Gln	Ser	Phe	Ser	Asn	
	460					465					470					
ctg	gga	ccg	aac	cgg	tgg	cac	aat	gaa	cgg	atc	acc	ttg	tca	act	ggc	1793
Leu	Gly	Pro	Asn	Arg	Trp	His	Asn	Glu	Arg	Ile	Thr	Leu	Ser	Thr	Gly	
	475				480					485					490	
aaa	agt	gta	tgg	ctt	ttc	aac	atc	aca	gcc	gac	cca	tat	gag	agg	gtg	1841
Lys	Ser	Val	Trp	Leu	Phe	Asn	Ile	Thr	Ala	Asp	Pro	Tyr	Glu	Arg	Val	
				495					500					505		
gac	cta	tct	aac	agg	tat	cca	gga	atc	gtg	aag	aag	ctc	cta	cgg	agg	1889
Asp	Leu	Ser	Asn	Arg	Tyr	Pro	Gly	Ile	Val	Lys	Lys	Leu	Leu	Arg	Arg	
			510					515					520			
ctc	tca	cag	ttc	aac	aaa	act	gca	gtg	ccg	gtc	agg	tat	ccc	ccc	aaa	1937
Leu	Ser	Gln	Phe	Asn	Lys	Thr	Ala	Val	Pro	Val	Arg	Tyr	Pro	Pro	Lys	
		525					530					535				
gac	ccc	aga	agt	aac	cct	agg	ctc	aat	gga	ggg	gtc	tgg	gga	cca	tgg	1985
Asp	Pro	Arg	Ser	Asn	Pro	Arg	Leu	Asn	Gly	Gly	Val	Trp	Gly	Pro	Trp	
	540					545					550					
tat	aaa	gag	gaa	acc	aag	aaa	aag	aag	cca	agc	aaa	aat	cag	gct	gag	2033
Tyr	Lys	Glu	Glu	Thr	Lys	Lys	Lys	Lys	Pro	Ser	Lys	Asn	Gln	Ala	Glu	
	555				560					565					570	
aaa	aag	caa	aag	aaa	agc	aaa	aaa	aag	aag	aag	aaa	cag	cag	aaa	gca	2081
Lys	Lys	Gln	Lys	Lys	Ser	Lys	Lys	Lys	Lys	Lys	Lys	Gln	Gln	Lys	Ala	
				575					580					585		

## 5800-79B SEQLIST.TXT

gtc tca ggt tca act tgc cat tca ggt gtt act tgt gga taa 2123  
 Val Ser Gly Ser Thr Cys His Ser Gly Val Thr Cys Gly \*

gcacaaatat ttctgttttg gttaaacttt aatcagttct tatctttcat ctgtttccta 2183  
 ggtaaaccag caaatitggc tcgataatat cgctggccta agcgtcaggc ttgttttcat 2243  
 gctgtgccac 2253

<210> 9  
 <211> 552  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Pfam consensus sequence for human sulfatase

<400> 9  
 Pro Asn Ile Leu Leu Ile Leu Ala Asp Asp Leu Gly Ile Gly Asp Leu  
 1 5 10 15  
 Gly Cys Tyr Gly Asn Pro Thr Ile Arg Thr Pro Asn Ile Asp Arg Leu  
 20 25 30  
 Ala Glu Glu Gly Leu Arg Phe Thr Asn Ala Tyr Val Thr Pro Leu  
 35 40 45  
 Cys Thr Pro Ser Arg Ala Ala Leu Leu Thr Gly Arg Tyr Pro His Arg  
 50 55 60  
 Thr Gly Met Tyr Thr Asn Asn Arg Ala Gly Val Leu Pro Phe Thr Gly  
 65 70 75 80  
 Trp Ser Leu Glu Gly Gly Leu Pro Leu Asp Glu Thr Thr Leu Pro Glu  
 85 90 95  
 Leu Leu Lys Glu Ala Gly Tyr Ala Thr Gly Met Val Gly Lys Trp His  
 100 105 110  
 Gly Tyr Asn Glu Glu Ser Ser Ala Ser Asp Phe Ala His Leu Pro Leu  
 115 120 125  
 Gly Arg Gly Phe Asp Tyr Phe Tyr Gly Asn Leu Gly Gly Glu Asp Gln  
 130 135 140  
 Trp Tyr Pro Leu Val Asp Ala Leu Leu Pro Phe Thr Asn Asp Thr Tyr  
 145 150 155 160  
 Thr Cys Glu Gly Gly Tyr Gly Phe Ser Lys Asp Val Ala Leu Lys Pro  
 165 170 175  
 Leu Gly Ala Leu Gly Val Asn Glu Val Glu Ala Pro Asp Lys Ala Leu  
 180 185 190  
 Ala Asp Tyr Lys Thr Ala Gly Ala Leu Asn Val Pro His His Val Phe  
 195 200 205  
 Glu Trp Ala Asp Arg Tyr Ala Gly Ala Val Asp Val Gly Arg Pro Phe  
 210 215 220  
 Leu Ala Val Leu Ile Phe Pro Arg Pro Ala Ala Cys Phe Leu Tyr Pro  
 225 230 235 240  
 Asn Ala Thr Val Val Ser Gln Pro Met Pro His Ser Pro Leu Thr Ala  
 245 250 255  
 Pro Arg Pro Trp Gln Leu Leu Ala Asp Glu Ala Leu Pro Phe Leu Glu  
 260 265 270  
 Arg Asn Gly Gln Arg Asp Lys Pro Phe Phe Leu Tyr Leu Ser Tyr Lys  
 275 280 285  
 His Val His Ile Pro Arg Asp Ala Pro Met Leu Phe Ser Ser Lys Asp  
 290 295 300  
 Phe Ala Gly Ser Ser Arg Arg Gly Leu Tyr Gly Leu Ile Leu Asp Ser  
 305 310 315 320  
 Val Glu Glu Met Asp Asp Gly Val Gly Arg Val Leu Asn Ala Leu Asp  
 325 330 335  
 Glu Leu Asn Gly Leu Leu Asp Asn Thr Leu Ile Ile Phe Thr Ser Leu  
 340 345 350  
 Leu Asp His Gly Gly His Leu Gly Ala His Gly His Leu Gly Ile Arg

## 5800-79B SEQLIST.TXT

```

      355      360      365
Ala Gly Gly Ser Asn Gly Pro Phe Arg Gly Gly Lys Gly Thr Asn Leu
      370      375      380
Tyr Glu Gly Gly Thr Arg Val Pro Leu Ile Val Arg Trp Pro Glu Gly
385      390      395      400
Ile Ile Ala Pro Gly Gln Val Ser Asp Glu Leu Val Ser Leu Met Asp
      405      410      415
Leu Phe Pro Thr Ile Leu Asp Leu Ala Gly Ala Pro Leu Pro Gly Val
      420      425      430
Ala Ala Gly Val Lys Asp Arg Ile Leu Asp Gly Val Ser Leu Leu Pro
      435      440      445
Leu Leu Leu Gly Ala Ala Gly Ser Ser Arg His Glu Thr Leu Phe Tyr
      450      455      460
Glu Ser Tyr Cys Asn Glu Gly Arg Gly Phe Leu Pro Ala Val Arg Trp
465      470      475      480
Gly Lys Lys Lys Ala His Phe Arg Thr Pro Asn Ile Ala Gly Trp Gln
      485      490      495
Arg Val Asp Phe Asp Asp Val Trp Lys Leu Phe Asn Thr Val Glu Asp
      500      505      510
Phe Asn Arg Ser Gly Asp Asp Ala Cys Arg His Gly Asp Val Cys Lys
      515      520      525
Cys Leu Gly Lys Pro Arg Arg Ser Val Thr His His Asp Pro Pro Leu
      530      535      540
Leu Tyr Asp Leu Ser Arg Asp Pro
545      550

```

&lt;210&gt; 10

&lt;211&gt; 520

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Pfam consensus sequence for human sulfatase

&lt;400&gt; 10

```

Pro Asn Val Leu Leu Ile Leu Ala Asp Asp Leu Gly Ile Gly Asp Leu
1      5      10
Gly Cys Tyr Gly His Pro Thr Ile Arg Thr Pro Asn Leu Asp Arg Leu
      20      25      30
Ala Glu Glu Gly Leu Arg Phe Thr Asn His Tyr Thr Ala Thr Pro Leu
      35      40      45
Cys Ser Pro Ser Arg Ala Ala Leu Leu Thr Gly Arg Tyr Pro His Arg
      50      55      60
His Gly Met Val Ser Asn Gly Arg Leu Gly Val Leu Gly Phe Thr Ala
65      70      75      80
Lys Ser Gly Gly Leu Pro Leu Asp Glu Thr Thr Leu Pro Glu Leu Leu
      85      90      95
Lys Glu Ala Gly Tyr Ala Thr Gly Leu Val Gly Lys Trp His Leu Gly
      100      105      110
Leu Asn Glu Asn Ser Asp Ala Ala Gly Asp Gly Glu His Leu Pro Leu
      115      120      125
Gly Trp Arg Gly Phe Asp Tyr Phe Asp Gly Phe Leu Tyr Gly Ser Pro
      130      135      140
Phe Thr Tyr Asp Glu Glu Asn Cys Asp Asn Gly Glu Gly Thr Glu Pro
145      150      155      160
Pro Glu Ala Tyr Pro Glu Gln Gly Trp Leu Pro Gln Ile Leu Gly Tyr
      165      170      175
Tyr Leu Thr Asp Leu Leu Ala Asp Lys Ala Leu Gly Leu Leu Asp Val
      180      185      190
Ala Ser Ala Ala Gly Arg Leu Leu Ala Lys Ala Leu Ala Ala Ser Arg
      195      200      205
Pro Phe Phe Leu Tyr Ile Ser Pro Pro Ala Pro His Phe Ser Ile Leu

```

## 5800-79B SEQLIST.TXT

210		215		220	
Phe Arg Asn Phe Lys Glu Val Ala Gln Pro Tyr Arg Ala Pro Gln Leu					
225 Thr Gln Leu Phe Val Asp Glu Ala Ala Asp Phe Ile Glu Arg Asn Lys					
Glu Lys Pro Phe Phe Leu Tyr Leu Ala Phe Leu Arg Leu His Val His					
Thr Pro Leu Phe Ser Pro Ala Glu Asp Leu Glu Ser Lys Asp Phe Leu					
Gly Arg Ser Gln Arg Gly Arg Tyr Gly Asp Leu Val Glu Glu Met Asp					
Asp Leu Val Gly Arg Val Leu Asp Ala Leu Glu Asp Leu Gly Leu Leu					
Asp Asn Thr Leu Val Ile Phe Thr Ser Asp Asn Gly Ala His Leu Glu					
Gly Thr Pro Glu Trp Tyr Gly Gly Gly Asn Gly Pro Leu Lys Gly Gly					
Lys Gly Tyr Gly Ser Leu Tyr Glu Gly Gly Ile Arg Val Pro Leu Leu					
Val Arg Trp Pro Gly Gly Ile Ala Pro Ala Gly Arg Val Lys Glu Lys					
Ser Glu Leu Val Ser His Val Asp Leu Ala Pro Thr Ile Leu Asp Leu					
Ala Gly Ala Pro Leu Pro Lys Val Ala Asn Gly Ala Lys Asp Arg Pro					
Leu Asp Gly Val Ser Leu Leu Pro Leu Leu Gly Gly Ala Ala Pro					
Ser Arg Arg Ala His Glu Thr Leu Phe His Tyr Asn Gly Lys Gly Arg					
Lys Leu Arg Ala Val Arg Trp Pro Arg Lys Ser Gly Lys Thr Pro Lys					
Leu Lys Ala His Phe Phe Thr Pro Ala Phe Asp Asp Thr Asn Asn					
Gly Trp Glu Cys Val Gly Thr Val Ser Gln Ala Asp Asp Ile Glu Asp					
Cys Arg Cys Glu Gly Val Glu Thr Val Thr His His Asp Pro Glu					
Leu Tyr Asp Leu Ser Arg Asp Pro					

<210> 11  
 <211> 1578  
 <212> DNA  
 <213> homo sapiens

<400> 11						
atgggctggc	tttttctaaa	ggttttgttg	gcgggagtga	gtttctcagg	atttctttat	60
cctcttgtgg	atttttgcac	cagtgggaaa	acaagaggac	agaagccaaa	ctttgtgatt	120
attttggccg	atgacatggg	gtggggtgac	ctgggagcaa	actgggcaga	aacaaaggac	180
actgccaaac	ttgataagat	ggcttcggag	ggaatgaggt	ttgtggattt	ccatgcagct	240
gcctccacct	gctcaccctc	ccgggcttcc	ttgtctaccg	gccggcttgg	ccttcgcaat	300
ggagtcacac	gcaactttgc	agtcacttct	gtgggaggcc	ttccgctcaa	cgagaccacc	360
ttggcagagg	tgctgcagca	ggcggtttac	gtcactggga	taataggcaa	atggcatctt	420
ggacaccaac	gctcttatca	ccccaacttc	cgtggttttg	attactactt	tggaatccca	480
tatagccatg	atatgggctg	tactgatact	ccaggctaca	accaccctcc	ttgtccagcg	540
tgtccacagg	gtgatggacc	atcaaggaac	cttcaaagag	actgttacac	tgacgtggcc	600
ctccctcttt	atgaaaacct	caacattgtg	gagcagccgg	tgaacttgag	cagccttgcc	660
cagaagtatg	ctgagaaagc	aaccacgttc	atccagcgtg	caagcaccag	cgggaggccc	720
ttcctgctct	atgtggctct	ggcccacatg	cacgtgccct	taccctgtgac	tcagctacca	780
gcagcgccac	ggggcagaag	cctgtatggg	gcagggctct	gggagatgga	cagtctggtg	840
ggccagatca	aggacaaagt	tgaccacaca	gtgaaggaaa	acacattcct	ctggtttaca	900
ggagacaatg	gcccgtgggc	tcagaagtgt	gagctagcgg	gcagtgtggg	tcccttcact	960

## 5800-79B SEQLIST.TXT

ggatttttggc	aaactcgtca	agggggaagt	ccagccaagc	agacgacctg	ggaaggaggg	1020
caccgggtcc	cagcactggc	ttactggcct	ggcagagttc	cagttaatgt	caccagcact	1080
gccttggttaa	gcgtgctgga	cattttttcca	actgtggtag	ccctggccca	ggccagctta	1140
cctcaaggac	ggcgctttga	tgggtgtggac	gtctccgagg	tgctcttttg	ccggtcacag	1200
cctgggcaca	gggtgctggt	ccaccccaac	agcggggcag	ctggagagtt	tggagccctg	1260
cagactgtcc	gcctggagcg	ttacaaggcc	ttctacatta	ccggtggagc	cagggcgtgt	1320
gatgggagca	cggggcctga	gctgcagcat	aagtttcctc	tgattttcaa	cctggaagac	1380
gataccgcag	aagctgtgcc	cctagaaaga	gggtggtgcg	agtaccaggc	tgtgctgccc	1440
gaggtcagaa	aggttcttgc	agacgtcctc	caagacattg	ccaacgacaa	catctccagc	1500
gcagattaca	ctcaggaccc	ttcagtaact	ccctgctgta	atccctacca	aattgcctgc	1560
cgctgtcaag	ccgcataa					1578

<210> 12  
 <211> 2616  
 <212> DNA  
 <213> homo sapiens

<400> 12						
atgaagtatt	cttgctgtgc	tctggttttg	gctgtcctgg	gcacagaatt	gctgggaagc	60
ctctgttcga	ctgtcagatc	cccagagttc	agaggacgga	tacagcagga	acgaaaaaac	120
atccgaccca	acattattct	tgtgcttacc	gatgatcaag	atgtggagct	ggggtccttg	180
caagtcattga	acaaaacgag	aaagattatg	gaacatgggg	gggccacctt	catcaatgcc	240
tttgtgacta	cacccatgtg	ctgcccgtca	cggtcctcca	tgctcaccgg	gaagtatgtg	300
cacaatcaca	atgtctacac	caacaacgag	aactgtctct	ccccctcgtg	gcaggccatg	360
catgagcctc	ggacttttgc	tgtatatctt	aacaacactg	gctacagaac	agcctttttt	420
ggaaaatacc	tcaatgaata	taatggcagc	tacatcccc	ctgggtggcg	agaatggctt	480
ggattaatca	agaattctcg	cttctataat	tacactgttt	gtcgcaatgg	catcaaagaa	540
aagcatggat	ttgattatgc	aaaggactac	ttcacagact	taatcactaa	cgagagcatt	600
aattacttta	aaatgtctaa	gagaatgtat	ccccataggc	ccgttatgat	ggtgatcagc	660
cacgctgcgc	cccacggccc	cgaggactca	gccccacagt	tttctaaact	gtaccccaat	720
gcttcccaac	acataactcc	tagttataac	tatgcaccaa	atatggataa	acactggatt	780
atgcagtaca	caggaccaat	gctgcccac	cacatggaat	ttacaaacat	tctacagcgc	840
aaaaggctcc	agactttgat	gtcagtggat	gattctgtgg	agaggctgta	taacatgctc	900
gtggagacgg	gggagctgga	gaatacttac	atcattttaca	ccgccgacca	tggttaccat	960
attgggcagt	ttggactggg	caaggggaaa	tccatgccat	atgactttga	tattcgtgtg	1020
ccttttttta	ttcgtgtgtc	aagtgtagaa	ccaggatcaa	tagtcccaca	gatcgtttct	1080
aacattgact	tggcccccac	gattcctggat	attgtctggg	tcgacacacc	tcctgatgtg	1140
gacggcaagt	ttgtcctcaa	acttctggac	ccagaaaagc	caggtaacag	gtttcgaaca	1200
aacaagaagg	ccaaaatttg	gcgtgatata	ttcctagtgg	aaagaggcaa	atttctacgt	1260
aagaaggaag	aatccagcaa	gaatatccaa	cagtcaaata	acttgcccaa	atatgaacgg	1320
gtcaaagaac	tatgccagca	ggccaggtac	cagacagcct	gtgaacaacc	ggggcagaag	1380
tggcaatgca	ttgaggatata	atctggcaag	cttcgaattc	acaagtgtaa	aggaccagt	1440
gacctgtctc	cagtccggca	gagcacgcgg	aacctctacg	ctcgcggctt	ccatgacaaa	1500
gacaaagagt	gcagtgttag	ggagtctggg	taccgtgcca	gcagaagcca	aagaaagagt	1560
caacggcaat	tcttgagaaa	ccaggggact	ccaaagtaca	agcccagatt	tgtccatact	1620
cggcagacac	gttcctttgc	cgtcgaattt	gaagggtgaa	tatatgacat	aaatctggaa	1680
gaagaagaag	aattgcaagt	gttgcaacca	agaaacattg	ctaagcgtca	tgatgaaggc	1740
cacaaggggc	caagagatct	ccaggtcttc	agtgggtggc	acaggggagc	gatgctggca	1800
gatagcagca	acgccgtggg	cccacctacc	actgtccgag	tgacacacaa	gtgtttttatt	1860
cttcccaatg	actctatcca	ttgtgagaga	gaactgtacc	aatcggccag	agcgtggaag	1920
gaccataagg	catacattga	caaagagatt	gaagctctgc	aagataaaaat	taagaatttta	1980
agagaagtga	gaggacatct	gaagagaagg	aagcctgagg	aatgtagctg	cagtaaacaa	2040
agctattaca	ataaagagaa	aggtgtaaaa	aagcaagaga	aattaaagag	ccatcttcac	2100
cgattcaagg	aggctgctca	ggaagtagat	agcaaactgc	aacttttcaa	ggagaacaac	2160
ccttagagga	agaaggagag	gaaggagaag	agacggcaga	ggaaggggga	agagtgcagc	2220
ctgcctggcc	tacttgctt	cacgcatgac	aacaaccact	ggcagacagc	cccgttcttg	2280
aacctgggat	ctttctgtgc	ttgcacgagt	tctaacaata	acacctactg	gtgtttgctg	2340
acagttaatg	agacgcataa	ttttcttttc	tgtgagtttg	ctactggctt	tttgaggtat	2400
tttgatatga	atacatatcc	ttatcagctc	acaataacag	tgacacaggt	agaacgaggc	2460
attttgaatc	agctacacgt	acaactaatg	gagctcagaa	gctgtcaagg	atataagcag	2520
tgcaacccaa	gacctaagaa	tcttgatgtt	ggaaataaag	atggaggaag	ctatgacctc	2580
cacagaggac	agttatggga	tggatgggaa	ggttaa			2616

## 5800-79B SEQLIST.TXT

<210> 13  
 <211> 1710  
 <212> DNA  
 <213> homo sapiens

<400> 13  
 atgcacaccc tctactggctt ctctctggtc agcctgctca gcttcggcta cctgtcctgg 60  
 gactggggcca agccgagctt cgtggccgac gggcccgggg aggctggcga gcagccctcg 120  
 gccgctccgc cccagcctcc ccacatcatc ttcatcctca cggacgacca aggctaccac 180  
 gacgtggggct accatgggtc agatatcgag acccctacgc tggacaggct ggcggccaag 240  
 ggggtcaagt tggagaatta ttacatccag cccatctgca cgccttcgcy gagccagctc 300  
 ctactggca ggtaccagat ccacacagga ctccagcatt ccatcatccg cccacagcag 360  
 cccaactgcc tgcccctgga ccaggtgaca ctgccacaga agctgcagga ggcaggttat 420  
 tccaccata tgggtgggcaa gtggcacctg ggcttctacc ggaaggagtg tctgcccacc 480  
 cgtcggggct tcgacacctt cctgggctcg ctcacgggca atgtggacta ttacacctat 540  
 gacaactgtg atggcccagg cgtgtgcggc ttcgacctgc acgagggtga gaatgtggcc 600  
 tgggggctca gcggccagta ctccactatg ctttacgccc agcgcgcccag ccatatcctg 660  
 gccagccaca gccctcagcg tccccctctt ctctatgtgg ccttccaggc agtacacaca 720  
 cccctcgagt cccctcgtga gtacctgtac cgctaccgca ccatgggcaa tgtggcccgg 780  
 cggaagtacg cggccatggg gacctgcatg gatgaggctg tgcgcaacat cacctgggcc 840  
 ctcaagcgct acggtttcta caacaacagt gtcatcatct tctccagtga caatggtggc 900  
 cagactttct cggggggcag caactggccg ctccgaggac gcaagggcac ttattgggaa 960  
 ggtggcgtgc ggggcctagg ctttgtccac agtcccctgc tcaagcgaaa gcaacggaca 1020  
 agccggggcac tgatgcacat cactgactgg taccggaccc tgggtgggtct ggcaggtggg 1080  
 accacctcag cagccgatgg gctagatggc tacgacgtgt ggccggccat cagcgagggc 1140  
 cgggcctcac cacgcacgga gatcctgcac aacattgacc cactctacaa ccatgcccag 1200  
 catggctccc tggaggggcg ctttggcatc tggaaacacc cgtgcaggc tgccatccgc 1260  
 gtgggtgagt ggaagctgct gacaggagac ccggtctatg gcgattggat cccaccgcag 1320  
 acactggcca ccttcccggg tagctgggtg aacctggaac gaatggccag tgtccgccag 1380  
 gccgtgtggc tcttcaacat cagtgtgac cttatgaac gggaggacct ggctggccag 1440  
 cggcctgatg tggctccgcac cctgctggct cgcctggccg aatataaccg cacagccatc 1500  
 ccggtacgct acccagctga gaacccccgg gctcatcctg actttaatgg ggggtgcttg 1560  
 gggccctggg ccagtgtatg ggaagaggag gaagaggaaag gaggggctcg aagcttctcc 1620  
 cggggtcgtc gcaagaaaaa atgcaagatt tgcaagcttc gatccctttt ccgtaaaactc 1680  
 aacaccaggc taatgtccca acggtatctga 1710

<210> 14  
 <211> 1800  
 <212> DNA  
 <213> homo sapiens

<400> 14  
 atggctccca ggggctgtgc ggggcatccg cctccgcctt ctccacaggc ctgtgtctgt 60  
 cctggaaaga tgctagcaat gggggcgctg gcaggattct ggatcctctg cctcctcact 120  
 tatggttacc tgcctgggg ccaggcctta gaagaggagt aagaaggggc cttactagct 180  
 caagctggag agaaactaga gccagcaca acttccacct cccagcccca tctcattttc 240  
 atcctagcgg atgatcaggg atttagagat gtgggttacc acggatctga gattaaaaca 300  
 cctactcttg acaagctcg tgccgaagga gttaaactgg agaactacta tgtccagcct 360  
 atttgcacac catccaggag tcagtttatt actggaaagt atcagatata caccggactt 420  
 caacattcta tcataagacc taccacaacc aactgtttac ctctggacaa tgccacccta 480  
 cctcagaaac tgaaggagggt tggatattca acgcatatgg tcggaaaatg gcacttgggt 540  
 ttttacagaa aagaatgcat gccaccaga agaggatttg atacctttt tggttccctt 600  
 ttgggaagtg gggattacta tacacactac aaatgtgaca gtcctgggat gtgtggctat 660  
 gacttgtatg aaaacgacaa tgctgcctgg gactatgaca atggcatata ctccacacag 720  
 atgtacactc agagagtaca gcaaacttta gcttccata accccacaaa gcctatatatt 780  
 ttatatattg cctatcaagc tgttcattca ccactgcaag ctctggcag gtatttcgaa 840  
 cactaccgat ccattatcaa cataaacagg aggagatatg ctgccatgct ttcctgctta 900  
 gatgaagcaa tcaacaacgt gacattggct ctaaagactt atggtttcta taacaacagc 960  
 attatcattt actcttcaga taatgggtggc cagcctacgg caggaggagg taactggcct 1020  
 ctcagaggta atattgggaa atattgggaa gggctgtagg ctttgtgcat ctttgtgcat 1080  
 agcccacttc tgaaaaacaa gggaacagtg tgtaaggaaac ttgtgcacat cactgactgg 1140  
 taccctactc tcatttctact ggctgaagga cagattgatg aggacattca actagatggc 1200  
 tatgatattc gggagaccat aagtgagggt cttcgctcac cccgagtaga tattttgcat 1260

5800-79B SEQLIST.TXT

aacattgacc	ccatatacac	caaggcaaaa	aatggctcct	gggcagcagg	ctatgggac	1320
tggaacactg	caatccagtc	agccatcaga	gtgcagcact	ggaaattgct	tacaggaaat	1380
cctgggtaca	gcgactgggt	ccccctcag	tctttcagca	acctgggacc	gaaccggtgg	1440
cacaatgaac	ggatcacctt	gtcaactggc	aaaagtgtat	ggcttttcaa	catcacagcc	1500
gacccatatg	agagggtgga	cctatctaac	aggtatccag	gaatcgtgaa	gaagctccta	1560
cggagggtct	cacagttcaa	caaaactgca	gtgccgggtca	ggtatcccc	caaagacccc	1620
agaagtaacc	ctaggctcaa	tggaggggtc	tggggaccat	ggtataaaga	ggaaaccaag	1680
aaaaagaagc	caagcaaaaa	tcaggctgag	aaaaagcaaa	agaaaagcaa	aaaaaagaag	1740
aagaaacagc	agaaagcagt	ctcaggttca	acttgccatt	caggtgttac	ttgtggataa	1800